DOCUMENT RESUME

ED 283 979 CE 047 647

AUTHOR Trace, Michael W.
TITLE Rehabilitation Associate Training for Employed Staff. Behavior Observation and Measurement (RA-5).
INSTITUTION Ellsworth Community Coll., Iowa Falls, IA.; Wisconsin Univ.-Stout, Menomonie. Stout Vocational Rehabilitation Inst.
SPONS AGENCY Rehabilitation Services Administration (ED), Washington, DC.
PUB DATE 82
GRANT 44-P-81418/7-02
NOTE 62p.; For related documents, see CE 047 642-655.
AVAILABLE FROM Materials Development Center, Stout Vocational Rehabilitation Institute, University of Wisconsin-Stout, Menomonie, WI 54751.
PUB TYPE Guides - Classroom Use - Guides (For Teachers) (052)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Behavioral Objectives; *Behavior Change; Charts; *Classroom Observation Techniques; *Counselor Training; Data Analysis; Data Interpretation; Evaluation Methods; Graphs; Inservice Education; Learning Activities; Learning Modules; Mathematical Formulas; *Measurement Techniques; Postsecondary Education; Records (Forms); Rehabilitation Counseling; *Vocational Evaluation; *Vocational Rehabilitation

ABSTRACT

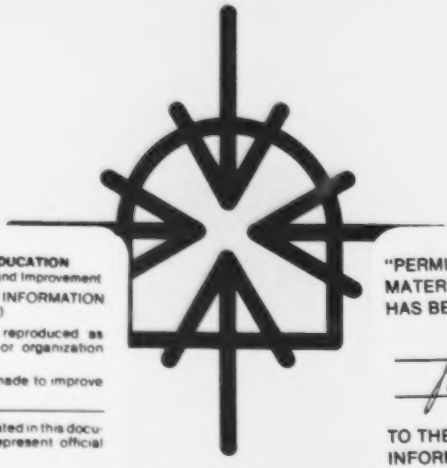
This learning module, which is intended for use in in-service training for vocational rehabilitation counselors, reviews the systematic methods available to measure and record client behavior. The first section covers procedures for making necessary calculations and keeping needed records with respect to behavioral events and their duration and interval. Also covered in the first section are procedures for use in time sampling, calculating a placheck (a term derived from the words "planned," "activity," and "check"), continuous recording, and developing permanent products. The second section deals with reliability for event and duration data and for interval and time sampling. Sample charts and graphs developed from various types of raw client observation data are presented. Five self-tests are also included. (MN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED283979

HUMAN SERVICES PROGRAM

ELLSWORTH COMMUNITY COLLEGE



U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

R. Fry

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

REHABILITATION ASSOCIATE
TRAINING FOR EMPLOYED STAFF

Behavior Observation And Measurement (RA-5)

CE047643

MDC

PRINTED AND DISTRIBUTED BY:
MATERIALS DEVELOPMENT CENTER

Stout Vocational Rehabilitation Institute
University of Wisconsin-Stout • Menomonie, WI 54751

BEST COPY AVAILABLE

Copyright, 1982, Merged Area (Education)
VI Community College District, an Iowa
school corporation

The Materials Development Center is partially
funded by the REHABILITATION SERVICES ADMINISTRATION,
U.S. DEPARTMENT OF EDUCATION, Washington, D.C.

BEHAVIOR OBSERVATION AND MEASUREMENT

by

Michael W. Trace, M.A.

Design: Debbie Anders-Bond, B.F.A.

Produced by
R.A.T.E.S.

Rehabilitation Associate Training
for Employed Staff

Ellsworth Community College
1100 College Avenue
Iowa Falls, Iowa 50126

Project Director: Michael J. Davis, Ph.D.

Editor: Michael W. Trace, M.A.

The production of this publication was supported in part by Grant No. 44-P-81418/7-02 from the Rehabilitation Services Administration. However, its contents are solely the responsibility of the authors and are in no way the responsibility of the Rehabilitation Services Administration.

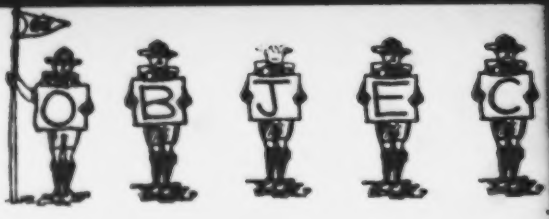
NOTE: Scattered throughout this module are questions which review the main points of the material presented. Be sure to answer these questions only in the student workbook!

REHABILITATION ASSOCIATE TRAINING
FOR EMPLOYED STAFF

MODULE: RA-5 Behavior Observation and Measurement

DESCRIPTORS: *Event recording, duration recording, interval recording, time sampling, placheck, permanent product, continuous recording, reliability, charting*

OVERVIEW: A critical element in all programming is the careful measurement and recording of client behavior. This module reviews the systematic methods available in doing this, the use and advantage of each, calculations to be made, and procedures for presenting data in chart form.



BEHAVIOR

1 Determine the most appropriate behavior recording operation to use depending upon the situation.

2 Calculate rate, % correct, placheck, percent of intervals, and % of samples, and reliability.

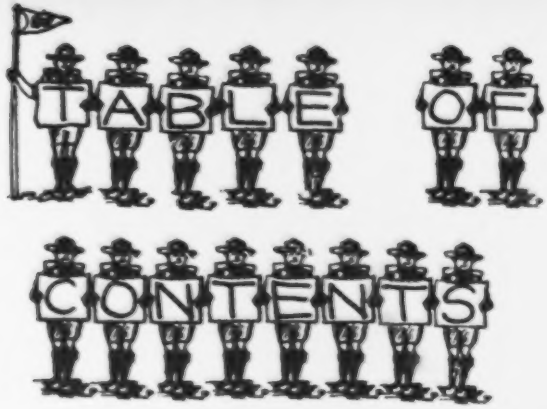
3 Chart behavior data, including appropriate labelling of vertical and horizontal axis, accurate plotting of data, and appropriate indication of treatment conditions.

EVALUATION

1 Class test. Will be given descriptions of situations and must list the appropriate operation to use.

2 Class test. Will be given hypothetical data

3 Class test. Will be given hypothetical data.



PAGE

Introduction----------8

Procedures----------10

Event Recording----------12
Calculations----------13

Duration Recording----------16
Calculations----------17

Interval Recording----------20
Calculations----------22

Time Sampling----------24
Calculations----------25

Placheck----------28
Calculations----------29

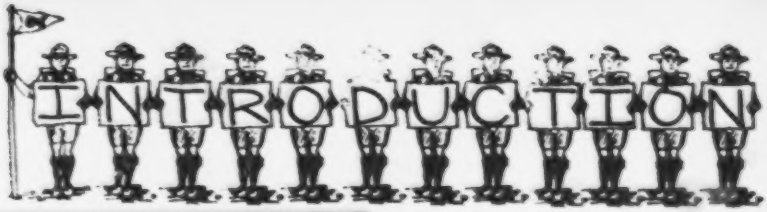
Continuous Recording----------30

Permanent Products----------33

Reliability----------36

For event and duration data----------39
For interval and time sampling----------40

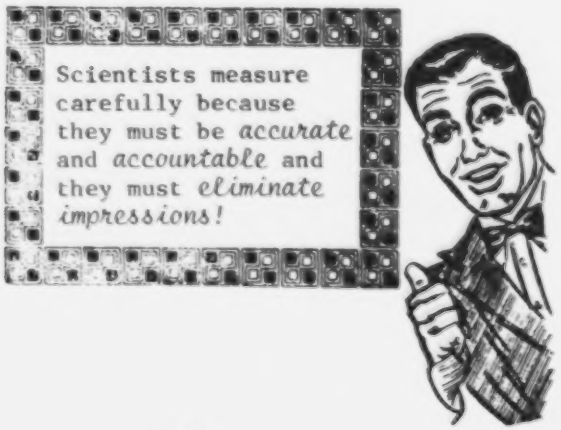
Records/Charts----------42



Scientists Measure Carefully

Walk into a physics or chemistry lab, and you will encounter a multitude of precise measuring instruments. These exist for a very specific purpose. In order to measure results of their activities, scientists must measure very accurately, and keep precise records in order to account for what they are doing. They cannot rely just on impressions they may have.

Scientists, therefore, measure very carefully. Quite often they are involved in counting *how many* or *how long;* they count some things and time the length of other things. When there are a large number to count all at once, they often take *samples*. Sometimes they do this recording *while* events occur; other times they do the recording *after* events have occurred (when there is something remaining to count or time). And often, for new events and situations that are not very clearly understood, they will write down almost *everything*. They will measure an event as well as what occurred before and after the event in order to determine what caused it.



Human Services is a Science

Dealing with human behavior is also a science, and like the other sciences, we must measure and record accurately. In order to account for interventions we are providing in clients' lives, we must be very clear on the effects our strategies are having on their behavior

because:

Simple impressions can be misleading (it may appear on the surface that clients are doing well).

We do not want to waste time implementing programs that are not needed (the behavior only appeared to be a problem).

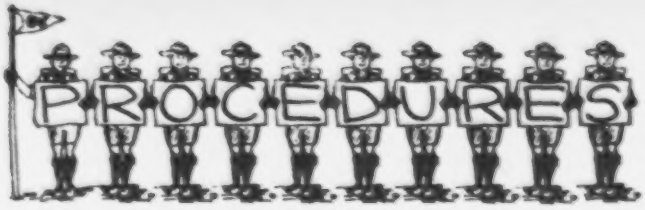
We do not want to waste time doing things that are not working (a particular training strategy is not working).

We must record accurately and account for what we are doing (to document for parents, guardians, accreditors, etc.).

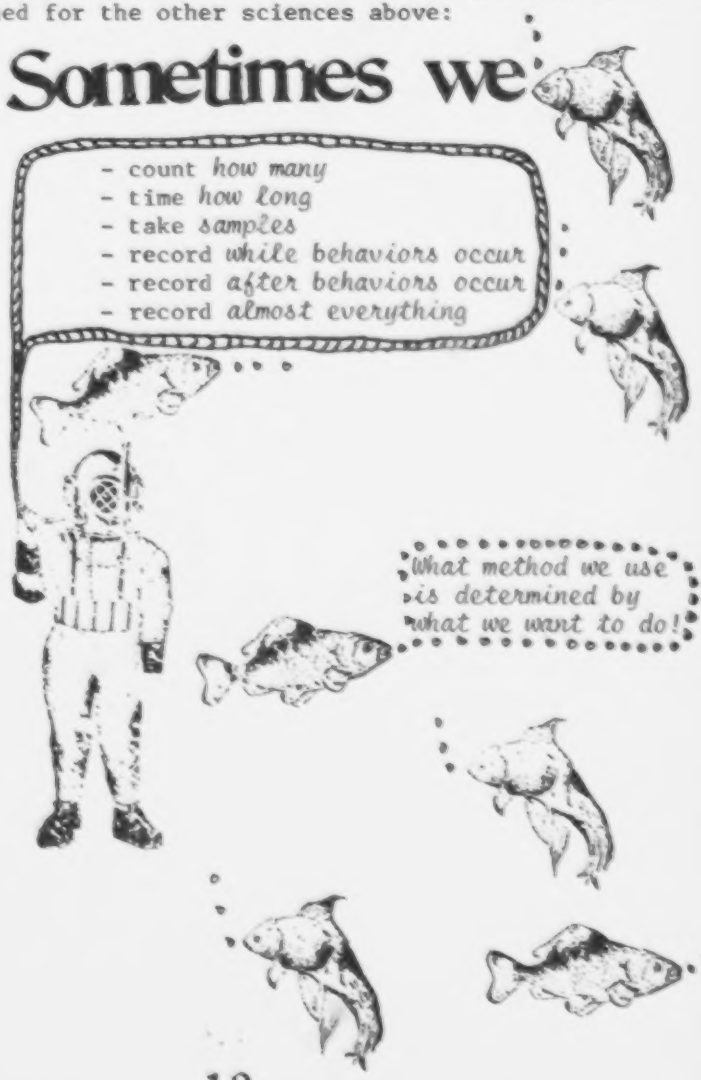
The slightest changes in behavior can indicate a needed program change. Often, if we can detect changes early, we can prevent a lot of problems and frustrations.

THE INFORMATION THAT FOLLOWS WILL PROVIDE YOU WITH SYSTEMATIC APPROACHES TO OBSERVING AND RECORDING BEHAVIOR IN HUMAN SERVICES.

Methods of observing & recording behavior



Measurement in human services is very much the same as we outlined for the other sciences above:

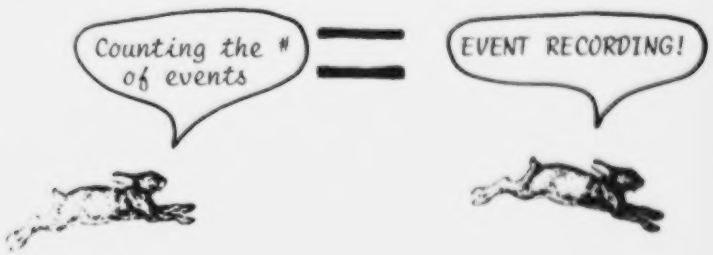

Sometimes we
- count how many
- time how long
- take samples
- record while behaviors occur
- record after behaviors occur
- record almost everything
What method we use is determined by what we want to do!

How do we know what method to use ?

HOW MANY?

(EVENT RECORDING)

When we are interested in *how many times* an event occurs (how many correct answers, how many bolts assembled, how many cards sorted, etc.) we do what is called *event recording*.



You will normally count the number of behaviors that occurred in a given period of time:

- *Number* of answers correct during training session.
- *Number* of bolts assembled in an hour.
- *Number* of times plate thrown on floor during meal time.

In order to count them, the events should have a clear beginning and end, and should be of moderate to low frequency. Events can be recorded:

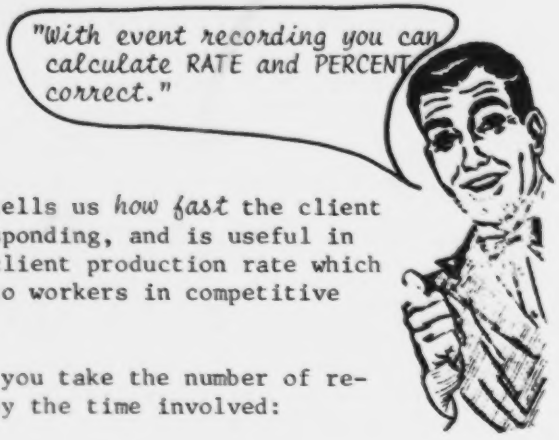
ON A CHECK LIST

WITH A HAND OR WRIST COUNTER

WITH PENCIL & PAPER

MOVING OBJECTS FROM ONE POCKET TO ANOTHER (tokens, beans, etc.)

CALCULATIONS



RATE → Rate tells us *how fast* the client is responding, and is useful in determining the client production rate which can be compared to workers in competitive jobs.

To compute rate, you take the number of responses divided by the time involved:

$$\frac{\text{\# of Responses}}{\text{Time}} = \text{RATE}$$

(*The formula* $\frac{R}{T}$ *is easy to remember because when you spell* R A T E, *the* R *comes before the* T.

EXAMPLES

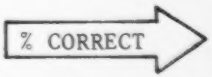
Here are some examples of computing rate:

DAY	# BOLTS ASSEMBLED	TIME	COMPUTATION		RATE
1	120	50'	$\frac{120\ (R)}{50\ (T)}$	=	2.4/min
2	30	40'	$\frac{30\ (R)}{40\ (T)}$	=	.74/min
3	680	2' 15"	$\frac{680\ (R)}{135\ (T)}$	=	5.03/min
4	150	70'	$\frac{150\ (R)}{70\ (T)}$	=	2.15/min

Note that sometimes:

R can be larger than T. (Days 1, 3, & 4)

R can be smaller than T. (Day 2)



We can also compute percent correct with event recording. This tells us what percent of all trials a client is correct on (eg., what percent of bolts were assembled correctly, what percent of holes were drilled correctly, etc.).

To compute percent correct, you divide the number of correct responses by the total number of responses and multiply times 100:

$$\frac{\#\ \text{Correct Responses}}{\text{Total}\ \#\ \text{Responses}} \times 100 = \%\ \text{Correct}$$

EXAMPLES

Here are some examples of computing % correct:

SESSION	# TRIALS	# CORRECT	COMPUTATION	% CORRECT
1	50	30	$\frac{30}{50}$ X 100 =	60%
2	140	110	$\frac{110}{140}$ X 100 =	79%
3	260	130	$\frac{130}{260}$ X 100 =	50%
4	80	80	$\frac{80}{80}$ X 100 =	100%

For our purposes here, whenever you are computing something and the answer is a PERCENT, you ALWAYS divide a smaller number by a larger number!

$\frac{\text{Smaller}}{\text{Larger}}$

HOW LONG?

(DURATION RECORDING)

Sometimes, *how long* a behavior occurs gives us a better measure. If John threw 5 tantrums today, and Bill threw only 2, you would think that John's behavior was much worse. However, if you knew John's 5 tantrums lasted a total of 50 seconds, and Bill's two tantrums lasted 40 minutes, you might think otherwise.

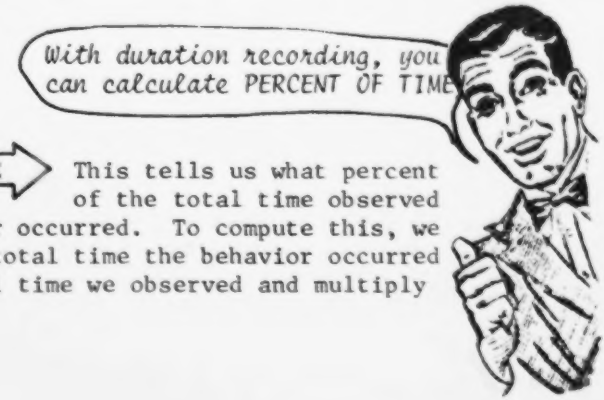
Timing how long a behavior occurs is known as *duration recording.*

Timing how long

= *Duration Recording*

A wall clock, wrist watch, stop watch, or time clock can be used for this.

CALCULATIONS



% OF TIME This tells us what percent of the total time observed the behavior occurred. To compute this, we divide the total time the behavior occurred by the total time we observed and multiply times 100:

$$\frac{\text{TIME BEHAVIOR OCCURRED}}{\text{TOTAL TIME OBSERVED}} \times 100 = \% \text{ of time.}$$

EXAMPLES

HERE ARE EXAMPLES OF COMPUTING PERCENT OF TIME:

DAY	TOTAL TIME OBSERVED	TIME BEHAVIOR OCCURRED	CALCULATION	% OF TIME
1	50'	25'	$\frac{25}{50}$ X 100 =	50%
2	2 hr	30'	$\frac{30}{120}$ X 100 =	25%
3	45'	45'	$\frac{45}{45}$ X 100 =	100%
4	1 hr 50'	1 hr 10'	$\frac{70}{110}$ X 100 =	64%

Let's see how we are doing. You now know that we can count *how many* (*EVENT RECORDING*) and *how long* (*DURATION RECORDING*). Data from these recording operations can give you:

Event Recording

RATE = $\frac{\text{\# of responses}}{\text{time}}$

PERCENT CORRECT = $\frac{\text{\# of correct trials}}{\text{total \# of trials}}$ X 100

Duration Recording

PERCENT OF TIME = $\frac{\text{time behavior occurred}}{\text{total time observed}}$ X 100

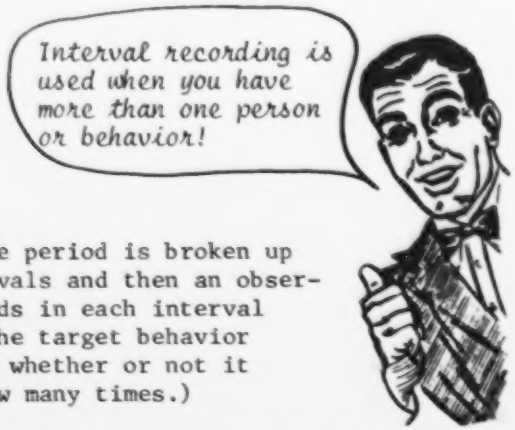
Turn to Self-Test #1 in the workbook to see how you can do in computing these!

Got it? You now are familiar with the *basics* of observing and recording behavior.

Unfortunately, these operations will not handle all situations. Can you imagine trying to do duration recording on crying behavior of five different individuals (timing all of them at once) or event recording for four different behaviors of a given individual (counting all of them at once)? This would be physically impossible. Also, it's difficult to count behaviors which occur at a very high rate.

INTERVAL RECORDING

All is not lost, however. Scientists have been very clever in coming up with reliable methods of observing more than one person or behavior. One of these is *interval recording*.



Basically, a time period is broken up into equal intervals and then an observer simply records in each interval whether or not the target behavior occurred. (Just whether or not it occurred, NOT how many times.)

The data sheet would be similar to the following, which has been divided into 10 second intervals:

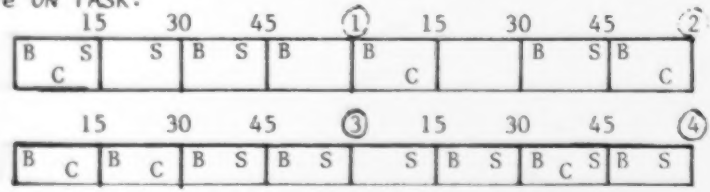
10	20	30	40	50	1	10	20

30	40	50	2	10	20	30	40

Using a stop watch, the recorder would observe the persons and/or behaviors during each interval and make a record if the behavior occurred (using initials, marks, etc). A record would be made in *every* interval the behavior occurred, even if it started in one interval and ended in the next (but it is only recorded *once* in each interval it occurred, no matter how many times it occurred in each interval).

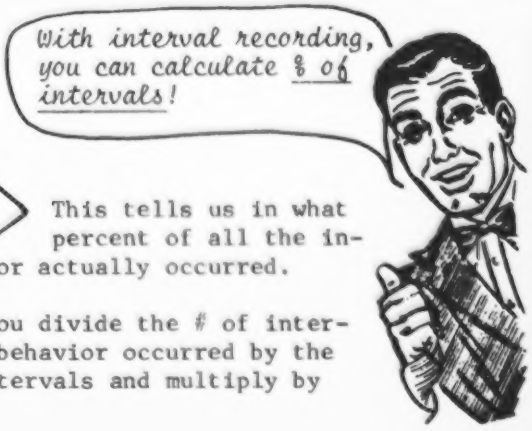
You can see how you can do more than one person or behavior this way. You do not have to count *how many*, just whether or not the behavior occurred. This *does*, however, give you representative data.

A completed data sheet would be similar to the following: Here we are observing 3 clients (B=Bob, C=Carol, S=Sue) and recording their initial in each interval that they were *ON TASK*:



(Naturally, some variations of this can occur. You may have their initials already listed in each interval, and then simply put a mark through the initial when the behavior occurs. Or, you may want to make a mark only if the behavior occurred during the *entire* interval, etc.

CALCULATIONS



% OF INTERVALS This tells us in what percent of all the intervals the behavior actually occurred.

To compute this, you divide the # of intervals in which the behavior occurred by the total number of intervals and multiply by 100.

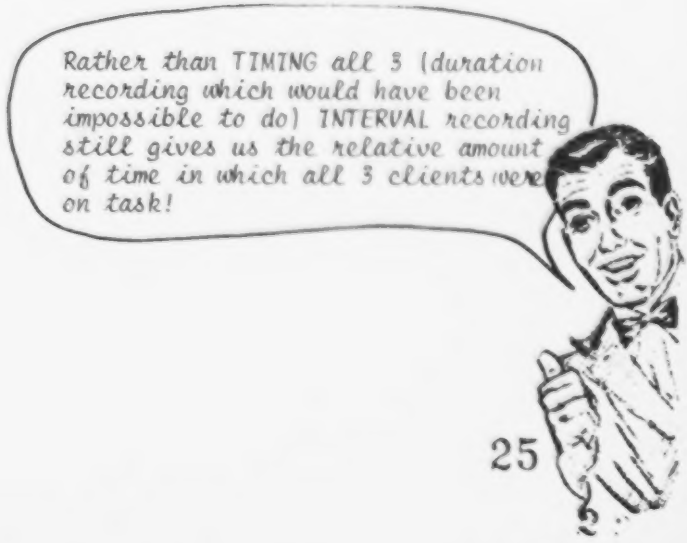
$$\frac{\text{\# INTERVALS BEHAVIOR OCCURRED}}{\text{TOTAL \# OF INTERVALS}} \times 100 = \text{\% of intervals}$$

For the data on the preceding page, Bob was on task for 13 intervals, Carol for 6 intervals, and Sue for 10 intervals. There were 16 intervals all together. Thus, computations for this data would be

EXAMPLES

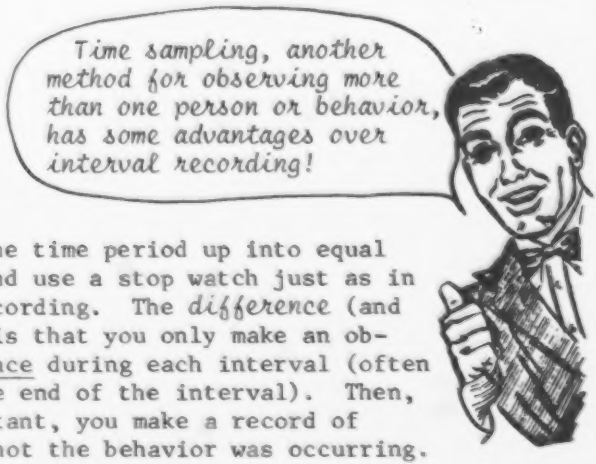
Here are examples of computing percent of intervals:

NAME	# INTERVALS BEHAVIOR OCCURRED	TOTAL # INTERVALS	COMPUTATION	% OF INTERVALS
Bob	13	16	$\frac{13}{16}$ X 100	81%
Carol	6	16	$\frac{6}{16}$ X 100	38%
Sue	10	16	$\frac{10}{16}$ X 100	63%


Rather than TIMING all 3 (duration recording which would have been impossible to do) INTERVAL recording still gives us the relative amount of time in which all 3 clients were on task!

TIME SAMPLING

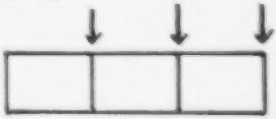
There is also one other method that can be used for observing more than one person or behavior, or a high rate behavior. Procedurally, it is very similar to interval recording.


Time sampling, another method for observing more than one person or behavior, has some advantages over interval recording!

You break the time period up into equal intervals and use a stop watch just as in interval recording. The *difference* (and advantage) is that you only make an observation once during each interval (often right at the end of the interval). Then, at that instant, you make a record of whether or not the behavior was occurring.

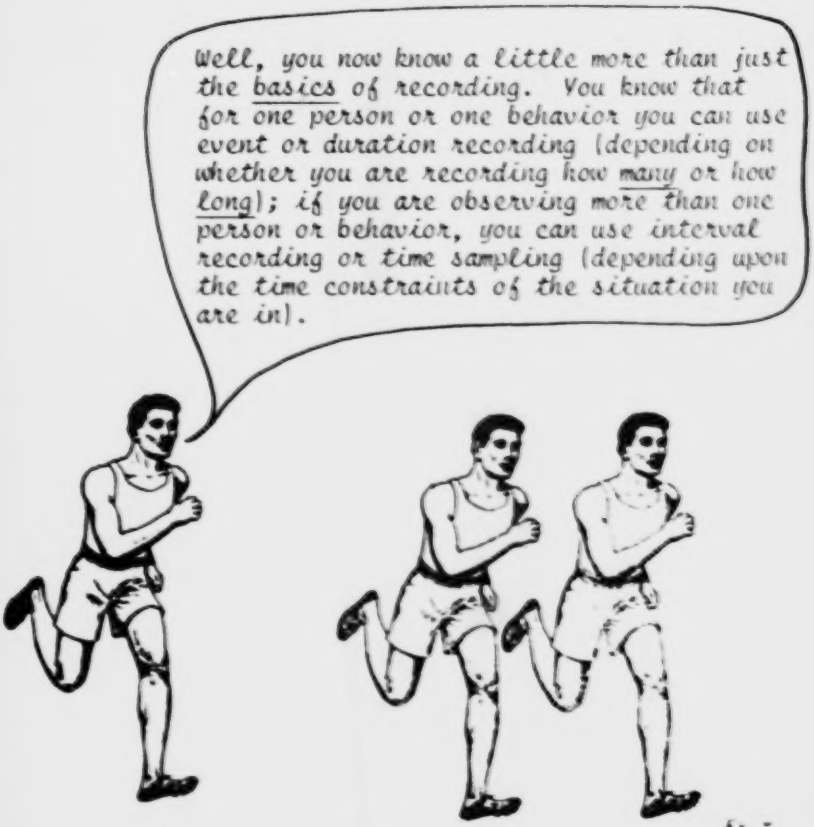
In interval recording you observe the entire interval!

In time sampling, you just observe at one instant (or sample) of the interval!



Time sampling is good if you are responsible for a large number of individuals (i.e., ten clients in a work activity center) but want specific data on a few of them. You only need to observe these few *once* during an interval; whereas in interval recording you need to observe them the entire interval, which does not leave time to watch the rest of the group.

That's the only DIFFERENCE with time sampling. The calculation procedures are exactly the same as interval recording, only your answers are called percent of SAMPLES rather than percent of intervals.


Well, you now know a little more than just the basics of recording. You know that for one person or one behavior you can use event or duration recording (depending on whether you are recording how many or how long); if you are observing more than one person or behavior, you can use interval recording or time sampling (depending upon the time constraints of the situation you are in).

23

If you look back over the material, you know how to calculate several operations:

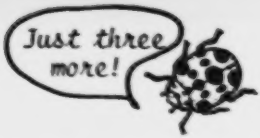
Rate

Percent Correct

Percent of Time

Percent of Intervals

Percent of Samples



We now are going to discuss just three more recording operations and you will be a master of the art. Given the discussion we had earlier, these are really very easy. They are just some procedures you can use for some specific situations.

For Group Behaviors

There are times when you are not really interested in the *specific* behaviors of *specific* individuals; but are more interested in measuring *group* behavior as a whole.

Does the presence of music in general help increase the on-task behavior of clients in a workshop?

What percent of clients in general are participating in the leisure activities I set up for them in the afternoon?

To make observations of *group* activity we do what is called

PLACHECK

CALCULATIONS

Procedurally, calculating PLACHECK is also somewhat similar to interval recording and time sampling. You break your time period up into intervals, and your data sheet would be similar to the following:

# observed												
# doing activity												

Then, you occasionally observe a sample of your group (a small part of the group) and record in the appropriate space the *number* you *observed*, and in the box below it the number who were actually *doing the activity*. After another interval of time, you observe another sample of the group and record your observations, then later another sample, etc., and your data would look something like this:

# observed	6	3	5	7	3	8	4	3	3	4	3	2	5
# doing activity	4	3	0	7	1	4	4	2	3	1	1	2	2

Two steps then remain:

Compute the actual percentage who were doing the activity for each of the observation samples (divide the smaller number by the larger number and multiply times 100)

Average the resultant percentages (add them up and divide by the number you have).

6	3	5	7	3	8	4	3	3	4	3	2	5	
4	3	0	7	1	4	4	2	3	1	1	2	2	
67%	100%	0%	100%	33%	50%	100%	67%	100%	25%	33%	100%	40%	= 815

$$\frac{815}{13} = 63\%$$

Thus in doing a planned activity check (PLACHECK) we find that on the average 63% are doing the activity.

Remember that PLACHECK is for *group* totals. We have not obtained any specific data on any one individual. This is what makes PLACHECK different from the earlier recording operations we have discussed.

NEW/UNCLEAR BEHAVIORS

Sometimes, it is difficult to work with clients because we are not clear about what the behaviors are. We may just need more information on the client's behaviors. It may be a new client whose social worker has simply told you, "he is an unreliable person." We do not really know what "unreliable person" means, or what the specific problem behaviors are.

In these situations, we can do what is called *continuous recording*. In general terms, continuous recording means that you observe the client and write down nearly everything that happens. In behavioral terms, we write down specific *behaviors* with their *antecedents* (what happened before the behavior) and *consequences* (what happened after the behavior).

We set up a sheet similar to the following:

ANTECEDENT (Before)	BEHAVIOR	CONSEQUENCE (After)

Then we make our observations, and can get an idea not only of what the specific behavior problems might be, but also what might be controlling the behaviors (what occurred before and after them).

Let's take a minute to see how this works!

Assume that Susan is a new client in a workshop and the following is a narrative of some behavior incidents:

> It's 8:00 and Susan's first day on the job. Her supervisor tells her to "*sort these IBM cards.*" Susan begins sorting the cards by number. After sorting a number of them Susan drops a few on the floor. Several clients laugh at her. Later, she drops some more and even more clients laugh. Soon she drops many. The supervisor returns and finds a mess of cards on the floor along with the fact that the cards were not sorted appropriately BY COLOR. The supervisor refers her to a counselor for "poor work behavior."

If one had been doing *continuous recording* during this period, the sheet would appear similar to the following:

ANTECEDENT (before)	BEHAVIOR	CONSEQUENCE (after)
1. Supervisor says, "*Sort these IBM cards.*"	2. Susan sorts by number.	
	3. Susan drops cards on the floor.	4. Clients laugh
	5. Susan drops more cards.	6. Clients laugh
7. Supervisor discovers mess plus cards in-appropriately sorted.	8. Susan goes to counselor	

This would be an example of continuous recording. More or less information might also be listed depending upon the specific observer.

With this type of data, here are some of the analyses that could be made of Susan's behavior:

1 Cards were sorted inappropriately by number instead of color (behavior #2) because the antecedent instructions were not clear (antecedent #1).

2 Susan's repeatedly dropping cards on the floor (behaviors #3 and #5) resulted in clients laughing at her (behaviors consequences #4 and #6) which appears to be increasing and thus reinforcing this behavior.

Thus, the supervisor needs to give clear instructions and arrange it so that Susan does not receive the attention for dropping cards on the floor. A simple analysis such as this and resultant changes may avoid a label of Susan as a "bad worker." Plus, she may now be learning that if certain behaviors exist, she can be sent to the counselor. If being sent to the counselor is reinforcing to her, these behaviors may increase.

THE POINT IS THAT CONTINUOUS RECORDING CAN HELP US ANALYZE BEHAVIORS AND CAN BE VERY USEFUL, PARTICULARLY FOR NEW CLIENTS.

And

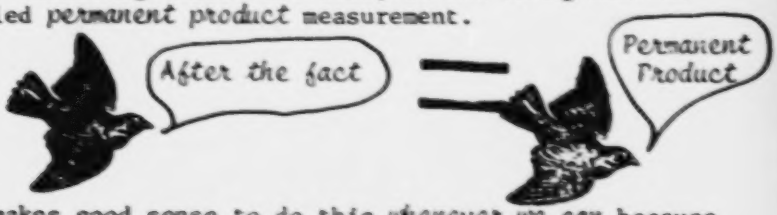
WE CAN SEE WHY IT IS IMPORTANT TO CAREFULLY OBSERVE AND RECORD BEHAVIOR. WE CAN OFTEN ELIMINATE PROBLEMS EARLY BEFORE BEHAVIORS BECOME STRONGER.

PERMANENT PRODUCTS

Now we come to our last recording method. Actually, it is more of a strategy than a method. Since all staff are usually very busy and do not always have time to stand and observe clients, as Marc Gold would say, *"there must be another way!"*

And for many behaviors there is. If John is assembling nuts and bolts, we do not have to stand there and count, we can come back after a period of time and *then* count how many he has assembled. The same thing is true with the IBM cards that we mentioned Susan was sorting.

MANY BEHAVIORS LEAVE A PRODUCT, SOMETHING THAT CAN BE TRACED AFTER THE FACT! In such instances, when we do our "measuring" after the fact, we are doing what is called *permanent product* measurement.



It makes good sense to do this *whenever we can* because we can measure a behavior quicker than standing and observing each instance as it occurs. You will find that this can be done with most production behaviors. The only time you would not necessarily want to do this would be to get a better idea of the specifics of the behavior. Then, you would use one of the other recording methods we talked about.

There you have it! Seven methods to observe and record behavior. Seems like it took quite a few pages to cover, and yet is really simple enough to summarize on this page:

METHOD	WHEN	WHAT
Event	Usually for one person/behavior of moderate frequency. Occurrences similar in length	Tells *how many*. Can compute *rate* & *percent correct*.
Duration	" "	Tells *how long*. Can compute *percent of time*.
Interval	For multiple clients behaviors & high frequency behaviors. Have time to observe just the target subjects.	Tells *percent of intervals* in which behavior occurred.
Time Sampling	Same as interval; can observe more than just target subjects.	Tells *percent of samples* in which behavior occurred.
Placheck	To measure planned activity of <u>groups</u>	Tells *overall percent of group* doing the activity.
Continuous	For new clients or unclear behaviors	Tells *antecedents, behaviors, & consequences*.
Permanent Product	After the fact. When behavior leaves a <u>trace</u>	Tells *how long, how many*, etc.

See Self-Test #3

When you go into a bank to cash a check, the teller will normally count the money once and then count the money in front of you. The purpose of the teller doing this is to provide a *double check* to ensure accuracy in handing out the money. Scientists also calibrate their instruments to insure they are accurate.

That is exactly what reliability is all about, ACCURACY. When you measure the behavior of clients, you want to be sure that your recording is accurate.

As good as we all are, certain problems can make our recording inaccurate:

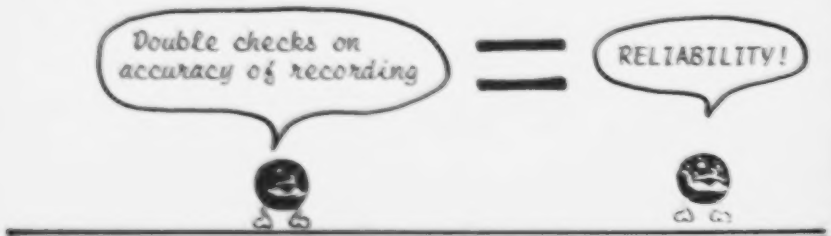
1. The definition of the behavior may not be clear. Thus, we may count it as "right" sometimes and "not right" other times.

2. Often, during the course of training, a definition may shift. We may observe that a client is trying very hard to do well. Without knowing it, we may "ease up" on our definition a little in order to encourage the client. Thus, improvements in behavior may be only due to the fact that our definition of the correct behavior has changed.

3. Sometimes, an observer's attention may drift away from the behavior being observed and some behaviors may be missed (the observer being distracted by a problem with another client, etc.).

4 Sometimes, an observer may just be *biased*. The client has not done well in other activities; the client is a "favorite" or "non-favorite." This also does not always occur purposely. Observers may unintentionally rate clients' behavior inaccurately as a result of liking the client or because of a client's attractiveness/unattractiveness or particular disability label.

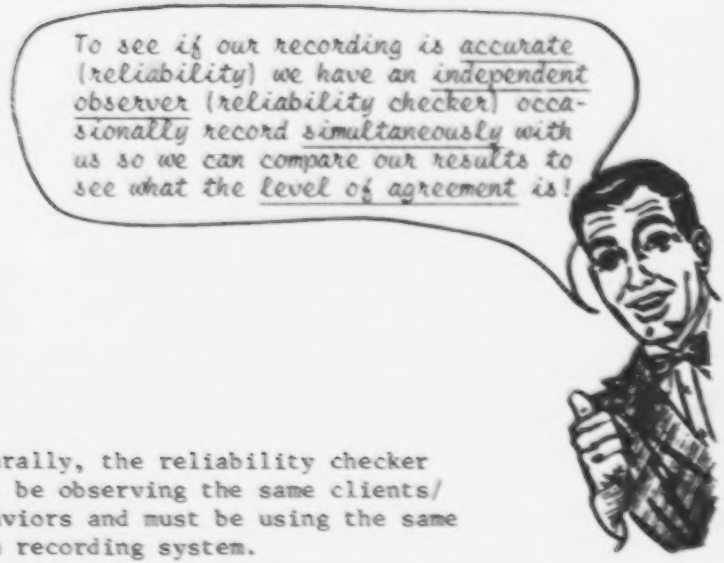
Thus, to ensure the accuracy of our recording, we sometimes take reliability checks. This is sort of a *double check* just as the teller does.



What we do is have another observer (reliability checker) also do the same recording we are doing. Then we can compare data for accuracy. Naturally, we will want to ensure that this observer is *independent* of us. Independence means the other observer:

- Cannot see our recording sheets
- Cannot see or hear us record
- Does not discuss data with us during the recording
- Can see the behavior clearly.

If the reliability checker is *not* independent (i.e., can hear the clicks of our stop watch, says in casual conversation, *"the client is not doing very well,"* etc.) then the results of our recording may be influenced by each other.


To see if our recording is accurate (reliability) we have an independent observer (reliability checker) occasionally record simultaneously with us so we can compare our results to see what the level of agreement is!

Naturally, the reliability checker must be observing the same clients/ behaviors and must be using the same data recording system.

Once both observers have collected their data, the level of agreement (percent) is calculated to determine reliability.

"FOR EVENT AND DURATION DATA"

Divide the smaller total by the larger total and multiply by 100.

$$\frac{\text{SMALLER}}{\text{LARGER}} \times 100$$

For example, if two of us are counting the number of times a worker walks away from his work station and. . .

I count: 37 times
Checker counts: 32 times

RELIABILITY would be $\frac{32}{37} \times 100$ or 86%

-OR-

If two of us are timing the total amount of time that an individual remains on task and. . .

I get: 2 hr 5' (125')
Checker gets: 2 hr 27' (147')

RELIABILITY would be $\frac{125}{147} \times 100$ or 85%

FOR INTERVAL & TIME SAMPLING

Determine the number of intervals the two observers *agreed* on that the behavior did or did not occur, and divide this by the total number of intervals and multiply times 100

$$\frac{\text{AGREED}}{\text{TOTAL}} \times 100$$

For example, assume we are doing interval or time sampling for on-task behavior of Bob, Carol, and Kevin. The following is a portion of the intervals for both the observer and the reliability checker. Indicated below each interval is whether the observers agreed (A) or disagreed (D) for each person.

	10			20			30			40			50		
Observer	B	C	K	B		K			K	B			B		
Rel. Checker	B		K			K	B	C	K	B	C	K		C		B
For Bob:	A			D			D			A			D		A
For Carol:		D			A			D			D			D		A
For Kevin:			A			A			A			D		A		A

If you count carefully the # of intervals they *agreed* upon for each person, you will get:

	# AGREED	COMPUTATION		RELIABILITY
BOB	3	$\frac{3}{6}$ X 100	=	50%
CAROL	2	$\frac{2}{6}$ X 100	=	33%
KEVIN	5	$\frac{5}{6}$ X 100	=	83%

40

What to accept

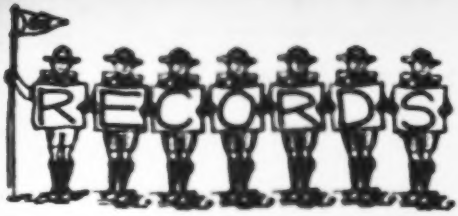
As a rule of thumb, if the reliability is 80% or higher, you can assume that the recording is accurate enough.

If it is less than 80%, you will want to investigate to see what the problem is (i.e., unclear definition of the behavior, the definition changed, etc.).

When to check

You don't need to check reliability every day. It's just a good idea to check it from time to time during programming. Initially, it is good to check it right after a new program begins, before losing too much time if the reliability does happen to be low. You may also check it again if treatment methods change or if you decide on a new definition of the behavior; you may check it if you notice some extreme changes in the client's behavior; and you should probably check right at the end of a program if you think the training has been successfully completed.

See Self-Test #4 to check yourself on the reliability section!

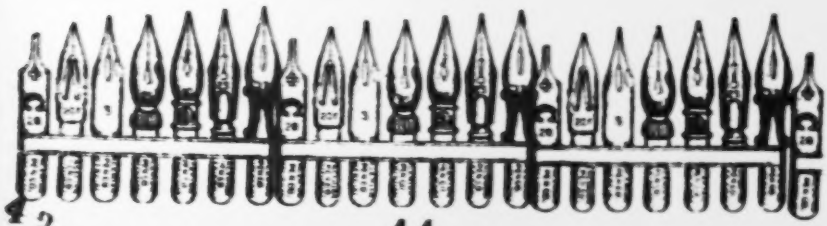


And

There are a multitude of methods for recording your data as you do your observations, and then summarizing it on graphs and charts.

RAW DATA

Raw data are the data and computations that you take down as you are actually doing the observations. How you do this is pretty much up to your imagination and organization. You can just record event and duration data on blank sheets of paper. For interval and time sampling you can use pre-designed sheets with intervals already drawn on them; you just need to write down the interval durations.



Event recording is often used with task analyses. The steps of the task are simply listed with spaces in which presence or absence can be checked off. Percent (of the total task) correct can also be computed if so desired.

In the example below, the task of shampooing hair can be easily measured by recording each sub-task as a separate event. If a more descriptive analysis is desired, the component tasks can be rated on a scale, for example,

1 = Independent
2 = Verbal prompt
3 = Model
4 = Physical guidance

Name;_____

Washing Hair

Date:					
1. Collect shampoo, towel, comb.					
2. Remove blouse or shirt.					
3. Place shampoo, towel, comb near sink.					
4. Adjust water to lukewarm					
5. Wet hair thoroughly					
6. Apply shampoo to hair					
etc.					

The check-list below shows a portion of a dining-skills chart. Breakfast, lunch, and dinner (B, L, and D) can be scored for each day. *For example*, 1 of the target behavior does not occur, 2 if it occurs sometimes but not throughout, and 3 if it continues throughout the entire meal.

	Monday			Tuesday			Wednesday			Thursday			Friday			Saturday		
	B	L	D	B	L	D	B	L	D	B	L	D	B	L	D	B	L	D
Eats bite-sized pieces (less than 1" square).																		
Uses correct utensils																		
Napkin is folded in half on lap.																		
Closes mouth while chewing.																		
Talks while mouth is free of food.																		

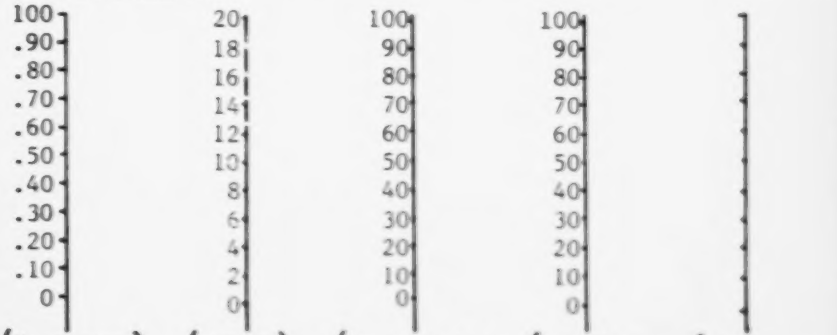
CHARTS

Most often, it is very helpful to translate your data into chart form. This helps provide a visual picture of the progress the client is making, and is much easier to read and interpret than a bunch of pages of raw data. Charts are extremely useful in showing other individuals (parents, advocates, clients themselves) the progress clients are making.

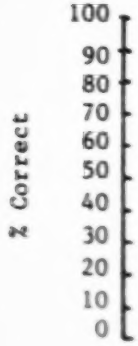
Think a minute about the products of the various recording operations we discussed:

Event Recording:	Number correct, rate, % correct
Duration recording:	Time, percent of time
Interval recording:	# of intervals, % of intervals
Time sampling:	# of samples, % of samples
Placheck:	% doing a planned activity
Permanent Product:	Could be any of the above

Scales for each of the products can be developed on a vertical line:

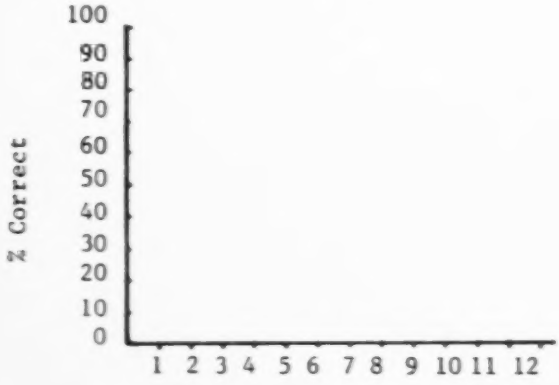


In order to make a chart, we can take our vertical line which has our measure on it (*for example, % correct*):



And

add a horizontal line to it on which we can plot respective training sessions, or days of training, etc:

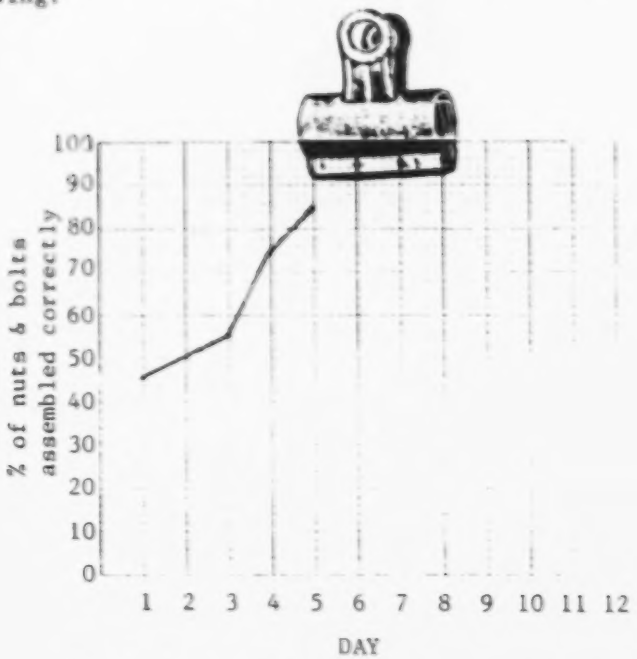


The result is a chart which is usually laid out on graph paper. We then can label the horizontal line (*with whatever our measure is* , and can label our vertical line (*with sessions, days, or whatever is applicable*). This allows us to make a mark within the graph on each respective day/session exactly what the client did on that day.

For example, let's say that a given client is assembling nuts and bolts, and each day we compute the percent that were assembled correctly. We might get the following:

Day 1	-	45%
Day 2	-	50%
Day 3	-	55%
Day 4	-	75%
Day 5	-	85%

If this were put on a chart, it would look like the following:

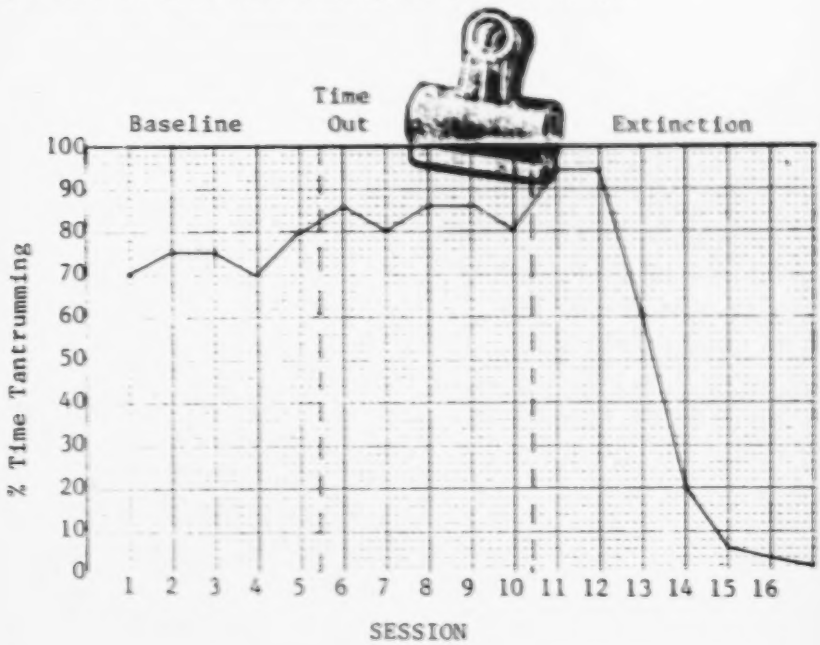


(*Each day, we place a dot at the interaction of that day and the % the client obtained that day*)

This can present a visual picture in an organized fashion which can show us a client's progress across an entire training program.

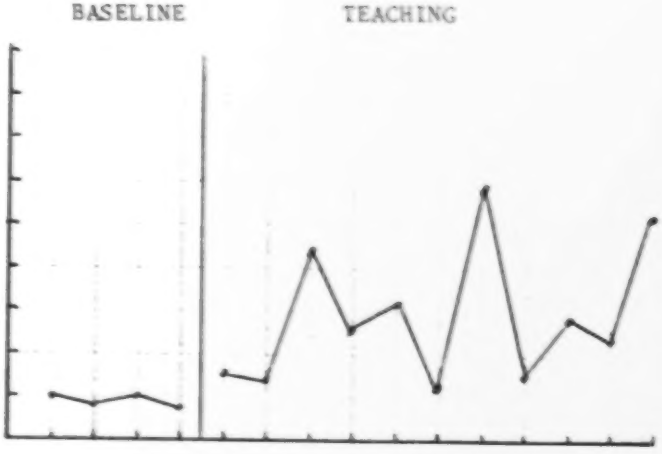
It is also helpful, sometimes, to draw dotted lines between respective days that training strategies changed, and then label on top of the chart what the respective strategies were.

In the chart below, we can see that percent of time tantrumming has been plotted for successive training sessions. Dotted lines separate the sessions, and the labels at the top of the graph indicate what was being done. First, a baseline was taken (that is the level of behavior before any training is attempted, so we can see how bad the behavior really is and can use this to see how effective our training is). Then, time out was attempted (which did not work, as the behavior did not change); and finally extinction was attempted (which did work as the percent of time tantrumming went down to 0).



Remember that data is to be used for *decision making*. You don't need to collect data on everything. . . just what can be used for decision making, accountability, etc. The following are examples of interpretations that can be made from data:

NON-FUNCTIONAL REINFORCER

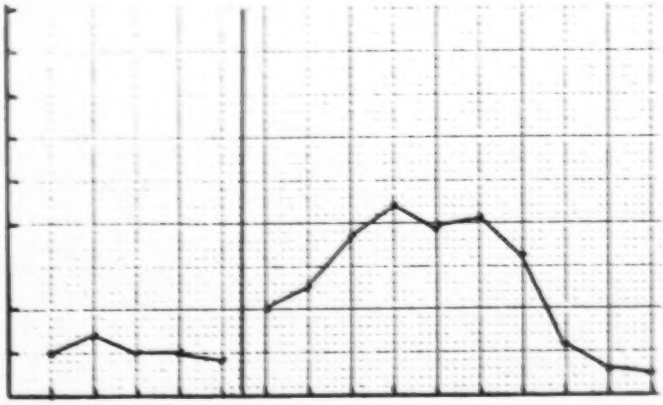


The current reinforcer is not effective every day, as indicated by the up and down pattern following baseline.

CHANGE REINFORCER OR ESTABLISH MENU

SATIATION

BASELINE TEACHING

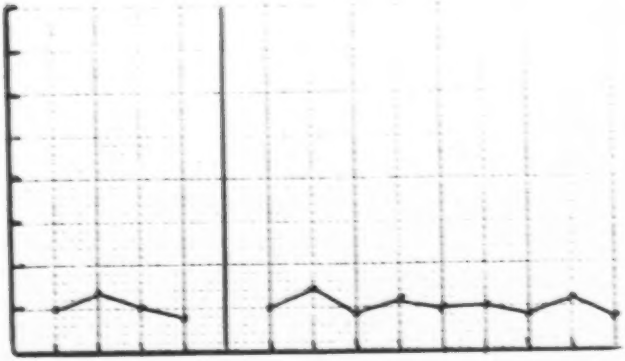


The reinforcer initially lead to an increase in behavior as indicated by the sharp rise following baseline, but is no longer effective as indicated by the subsequent downward trend.

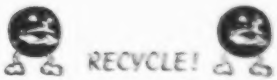
CHANGE REINFORCER OR ESTABLISH MENU

PROGRAM INAPPROPRIATE

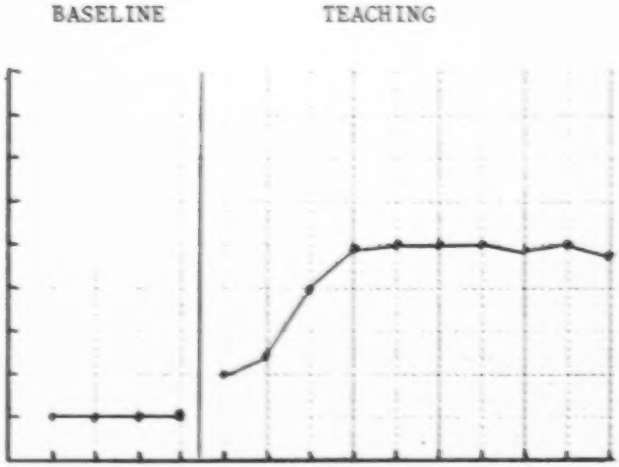
BASELINE TEACHING



Teaching is providing no change in behavior.



REACHED A STEP THAT IS TOO HARD



After a sharp increase, progress has leveled off

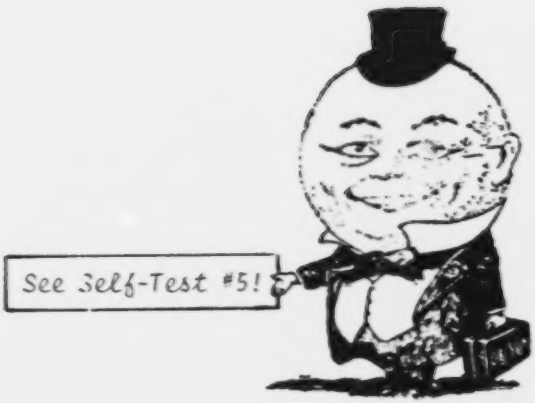
RE-DO CONTENT FOR THAT STEP

That gives you the basics of charting behavior. The main thing to keep in mind is to be *very neat*:

Use a ruler and draw straight lines

Carefully label your horizontal and vertical lines, and your treatment conditions.

Plot your data points very carefully. On a given day or session, in addition to plotting your observation, you can place an asterisk (*) to indicate what your reliability checker observed on days reliability is taken.


See Self-Test #5!

That also ends this module. Hope you now now can:

Observe and record accurately

Make necessary computations

Plot these on charts so others can see the results!

Self-Test #1

1. Calculate the following as indicated:

a) Calculate rate

Day	# Assembled	Time	Rate
1	150	30 min.	
2	150	3 hrs.	
3	60	2 hrs/10 min.	
4	340	1 hr/5 min.	
5	100	1 hr/40 min.	

b) Calculate % correct

Session	Trials	% Correct	% Correct
1	50	25	
2	80	80	
3	130	20	
4	150	110	
5	105	0	

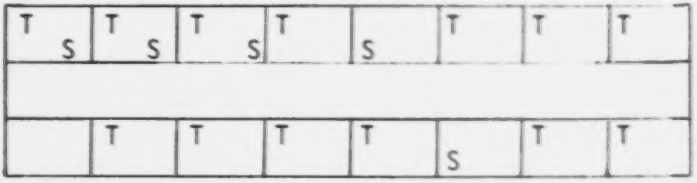
c) Calculate % of time

Day	Time Observed	Time Crying	% Time Crying
1	2 hrs.	1 hr/10 min.	
2	3 hrs/10 min.	40 min.	
3	70 min.	20 min.	
4	45 min.	45 min.	
5	4 hrs.	2 hrs/30 min.	

Self-Test #2

1. Compute percent of intervals for each behavior:

T=tantrums S=spitting



T= S=

2. Compute the above assuming it was time sampling.

T=
S=

Self-Test #3

1. List the appropriate behavior recording technique for each of the situations below:

   a) I'm going to supervise a client they say is very "distracting" but I do not know what the specific behaviors are.

   b) I want to determine how many times John runs to the bathroom each day and what this averages per hour.

   c) You are a supervisor for ten clients in a workshop and would like to determine how much four of them work.

   d) In situation "c" you get another staff person to supervise the six clients you are not observing.

   e) You are interested in the times it takes Barb to get to her work station each morning after walking into the workshop.

   f) You want to determine if music has an overall effect on clients remaining at their work station.

   g) You want to determine at the end of the day how many pallets several clients have assembled.

Self-Test #4

Compute reliability in each of the following examples:

1. Time-sampling.
   F-fighting

	Interval											
	1	2	3	4	5	6	7	8	9	10	11	12
Observer 1	F	F		F		F		F	F	F		F
Observer 2		F		F	F		F		F	F		

Percent agreement = _____

2. Duration--time required to complete chore.

   Observer 1 35 minutes

   Reliability observer 30 minutes

   Percent of agreement = _____

3. What is one reason why reliability might be too low?

Self-Test #5

1. Chart the following data. Be sure to label all conditions and axes.

Day	% of bolts assembled correctly	
1	35%	Baseline
2	30%	
3	28%	
4	32%	
5	34%	
6	70%	With use of model
7	85%	
8	80%	
9	83%	
10	85%	
11	85%	